DESIGN OF EXPERIMENTS (DOE)
HISTORY

- Developed in the early 1920s by Sir Ronald Fisher at the Rothamsted Agricultural Field Research Station in London, England.
- His initial experiments were concerned with determining the effect of various fertilizers on different plots of land.
- The final condition of the crop was not only dependent on the fertilizer but also on a number of other factors (such as underlying soil condition, moisture content of the soil, etc.) of each of the respective plots.
- Fisher used DOE which could differentiate the effect of fertilizer and the effect of other factors.
- Since then DOE has been widely accepted and applied in biological and agricultural fields as well as industry.

Examples

Medicine:

- Does salted drinking water affect blood pressure (BP) in mice?

  1% NaCl water 14 days Measure BP

- How does a company test a new vaccine/drug?

Industry:

How does a manager check the performance of machines or operators in a factory?

Rotate the machines, workers, shifts, days

Agriculture:

How does one find out the effect of fertilizer on crops?

SOME DEFINITIONS IN DOE

- Experiment: An investigation in which the investigator applies some treatments to experimental units and then observes the effect of the treatments on the experimental units by measuring one or more response variables.
- Treatment Factors and their levels:
  - *Treatment factor* is used to mean any substance or item whose effect on the data is to be studied.
  - *The levels* are the specific types or amounts of the treatment factor that will actually be used in the experiment.
  - For example, a treatment factor might be a drug or a chemical additive or temperature or teaching method, etc. The levels of such treatment factors might be the different amounts of the drug to be studied, different types of chemical additives to be considered, selected temperature settings in the range of interest, different teaching methods to be compared, etc.
  - If the levels of a treatment factor are quantitative (i.e., can be measured), then they are usually chosen to be equally spaced.
  - For convenience, treatment factor levels can be coded.
  - We will sometimes use the term treatment to mean a level of a treatment factor in a single factor experiment, or to mean a treatment combination in a factorial experiment.
- Variable: All the factors, their levels and all the measured traits (responses) are variables. They are attributes, characteristics, or dimensions being measured or studied.

SOME DEFINITIONS IN DOE

- Cases:The objects being studied, also known as subjects or participants.
- Independent or class variables: Factors applied to the experimental units are independent variables. Other names include predictor variables, regressors, controlled variables, manipulated variables, explanatory variables, or input variables.
- Response or dependent variables: Measured or observed traits in the experiment. The strength of the paper measured after treated with starch is a response (or dependent) variable. Other names include response variable, the regressand, the measured variable, the responding variable, the explained variable, the outcome variable, the experimental variable or the output variable.
- Grouping: a.k.a subject variable, a type of independent variable with naturally occurring groups such as males and females.
- Main effects: This is the simple effect of a factor on a dependent variable. It is the effect of the factor alone averaged across the levels of other factors. Example: In the starch experiment, it was found that on average, the starch increased the strength (main effect) compared to the control (no starch).

SOME DEFINITIONS IN DOE

- Interaction effect: An interaction is the variation among the differences between means for different levels of one factor over different levels of another factor. Example: Lets say starch and latex both significantly increased the paper strength (main effects). However, when starch and latex applied to the same experimental unit, it was found that the increase in paper strength was even higher. The paper got the benefits of both factors, plus a bonus.
- Experimental (or Sampling) Units: They are the "material" to which the levels of the treatment factor(s) are applied. For example, in agriculture these would be individual plots of land, in medicine they would be human or animal subjects, in industry they might be batches of raw material, factory workers, etc. If an experiment has to be run over a period of time, with the observations being collected sequentially, then the times of day can also be regarded as experimental units. Exptal units must be representative.

SOME DEFINITIONS IN DOE

- Population: A population is any entire collection of people, animals, plants from which we may collect data. i.e. the entire collection of items which is the focus of concern. It is the entire group we are interested in, which we wish to describe or draw conclusions about. Example, we are interested in average GPA values of all KNUST students. All the students registered in KNUST is a population. A population can be of any size. The members of the population must share at least one measurable feature.
- Sample: A sample is a group of units selected from a larger group (the population). By studying the sample it is hoped to draw valid conclusions about the larger group. A sample is generally selected for study because the population is too large to study in its entirety.

NOTE!!!The critical difference between a population and a sample, is that with a population our interest is to identify its characteristics whereas with a sample, our interest is to make inferences about the characteristics of the population from which the sample was drawn.

SOME DEFINITIONS IN DOE

- Parameter: A parameter is a value, usually unknown, and which therefore has to be estimated. It is used to represent a certain population characteristic. For example, the population mean is a parameter that is often used to indicate the average value of a quantity. Parameters are often assigned Greek letters (e.g. σ for standard deviation).
- Statistic: A statistic is a quantity that is calculated from a sample of data. It is a measure of the items in a random sample. For example, the mean, a variance and a standard deviation calculated from a sample are statistics. The only reason to ever draw a random sample is to infer something about the population from which it came. When we calculate a given statistic we only do so in order to estimate a corresponding parameter of the population from which the sample was drawn.Statistics are assigned Roman letters (e.g. s for standard deviation). The value of a statistic will vary from sample to sample.

SOME DEFINITIONS IN DOE

- Estimate: An estimate is an indication of the value of an unknown quantity based on observed data. Example, suppose we want to know the average height of KNUST students (population). We can use an estimate of this population mean by calculating the mean of a sample of students. Estimators of population parameters are sometimes distinguished from the true value by using the symbol 'hat'. For example, σ =true population standard deviation. $\hat{\sigma}$ estimated (from a sample) population standard deviation.
- Sampling Distribution: The sampling distribution is the probability distribution or probability density function of the statistic.
  - The Sampling Distribution of a statistic is the set of values that we would obtain if we drew an infinite number of random samples from a given population and calculated the statistic on each sample.
  - Example: Throwing enough darts at a dart board will soon form a pattern that would then represent the distribution of the process of that particular dart thrower throwing darts at that dart board from a fixed distance.
  - Derivation of the sampling distribution is the first step in calculating a confidence interval or carrying out a hypothesis test for a parameter. Example, suppose that $x_1, x_2, \ldots, x_n$ are a simple random sample from a normally distributed population with expected value and known variance s_2. Then the sample mean x is normally distributed with expected value μ and variance s_2/n.

Three Fundamental Experimental Design Concepts Attributed to R.A. Fisher

- Randomization: random assignment of treatments to experimental units.

- Blocking: grouping similar experimental units together and assigning different treatments within such groups of experimental units. These are 2 examples:
  - A metallurgist wants to improve the strength of a steel product. Four factors are being considered for the experiment, which might have some impact on the strength. It is decided to study each factor at 2-levels (i.e. a low setting and a high setting). An eight trial experiment is chosen by the experimenter but only four trials are possible to run for each day. Here each day can be treated as a separate block.
  - An experiment in a chemical process requires 2 batches of raw materials for conducting the entire experimental runs. In order to minimize the batch-to-batch variability, we need to treat each batch of raw material as a noise factor. In other words, each batch of raw materials would form a block.

- Replication: applying a treatment independently to two or more experimental units.
  - There is a difference between replication and repeated measurements. For example, suppose four subjects are each assigned to a drug and a measurement is taken on each subject. The result is four independent observations on the drug. This is replication.
  - On the other hand, if one subject is assigned to a drug and then measured four times, the measurements are not independent. We call them repeated measurements.

ART OF EXPERIMENTATION

Experimenter's Initial Questions

- How many observations do I need to take?
- Given my limited budget, how can I gain as much information as possible?
- What is the main purpose of running this experiment?
- What do I hope to be able to show?

Reasons for running an Experiment

- to determine the principal causes of variation in a measured response,
- to find the conditions that give rise to a maximum or minimum response,
- to compare the responses achieved at different settings of controllable variables,
- to obtain a mathematical model in order to predict future responses.

Observational and Experimental Studies

- Observations can be collected from observational studies as well as from experiments, but only an experiment allows conclusions to be drawn about cause and effect.
- Research studies may often be classified as either observational or experimental, although some are a mixture of the two.
- In an observational study, data are collected without any attempt to manipulate or influence the outcome. For example:
  - Fish may be collected from three different regions of a lake, in order to compare their weights over the three locations.
  - Children from three different schools may be compared for their performance on an achievement test.
  - Households from three suburbs are surveyed to compare their incomes and political opinions.
  - People with a particular disease are studied, to see if the severity of their symptoms is related to their age, smoking habits, diet or amount of exercise.

Observational and Experimental Studies

- In experiments usually some manipulation is attempted, in order to see if the outcome is related to the factor being controlled. For example:
  - Twenty plots of carrots are grown in a field. Each plot is randomly allocated to one of five fertilizers, with four plots for each fertilizer. At the end of the experiment, the carrots from each plot are weighed. The yield of carrots with different fertilizers is being studied.
  - Twenty children from a class are each randomly assigned to one of five different teaching methods, four children to each method. After three weeks of teaching, each child is tested for understanding of the material taught. The different teaching methods are being compared.
  - Eggs are incubated in the laboratory at chosen temperatures, with the outcome of interest being the sex of the hatchling.
  - Fifty subjects are assigned randomly to one of five different working conditions (quiet, loud pop music, people talking, etc.) There are ten in each group. They are then tested repeatedly to find each persons average response time to a particular type of stimulus (e.g. hearing a fire alarm bell). The influence of background noise on response time is being investigated.
  - People with a certain disease are randomly allocated to three different drugs. The drugs are being compared for their influence on the progress of the disease.

Designing/ Planning an Experiment

The following steps summarize a very large number of decisions that need to be made at each stage of the experimental planning process. The steps are not independent, and at any stage, it may be necessary to go back and revise some of the decisions made at an earlier stage.

- Define the objectives of the experiment.
- Identify all sources of variation, including:
  - treatment factors and their levels,
  - experimental units,
  - blocking factors, noise factors, and covariates.
- Choose a rule for assigning the experimental units to the treatments.
- Specify the measurements to be made, the experimental procedure, and the anticipated difficulties.
- Run a pilot experiment.
- Specify the model.
- Outline the analysis.
- Calculate the number of observations that need to be taken.
- Review the above decisions. Revise, if necessary.

IMPORTANT !!! It's only after all of these decisions have been made should the data be collected and analyzed and a report written.

KEY POINTS TO NOTE

A designed experiment is a methodical way to conduct a series of trials by optimizing the means at the disposal of the researcher: time and money. But this method is not a substitute for actual experiments, and especially cannot replace human intelligence and ingenuity. It is the experimenter that has to

- **Define the objective** It essential to precisely quantify the objective of the study.
- **Pose the problem well** A poorly posed problem cannot be saved by a designed experiment.
- **Be imaginative and creative**
  Conducting a designed experiment is not only a technique. It helps with the interpretation of the results, but it is the experimenter who is responsible for the hypotheses and conclusions.
- **Choose the factors**
  If the experimenter forgets an important factor of the study, it is he who is responsible for anomalies and erroneous conclusions.

KEY POINTS TO NOTE

- **Define limits on the study domain**
  This is one of the big difficulties encountered by the experimenter. If the domain is too small, the hoped-for responses may not be inside the domain, which means additional trials. If the domain is too large, the postulated model will not serve well in interpretation of the results.
- **Carefully conduct the experiments**
  Badly conceived experiments, uncontrolled factors, vague measurements, errors in notation, location, or transcription cannot be repaired by an experimental design. On the contrary: they require a great deal of care in the experiments themselves to pull out all the information.
- **Choose the responses**
  This is an equally vital point for the quality of interpretation. All responses are not equal. Some are not precise enough; others do not provide effects that are sufficiently clear-cut to draw conclusions from. Finally, it's necessary that the chosen responses allow us to answer the question of the study.
  In summary, the designed experiment is a useful and strong technique, but both control of the experiment and the responsibility for the conclusions drawn remains firmly in the hands of the experimenter.

STANDARD DESIGNS

Factorial Designs

2-level Factorial Designs

- Each experimental variable (factor, k) has 2 levels
- Each combination is tested
- Number of combinations is 2^k
  $\longrightarrow$ 2^k factorial design.
- n is a power of 2.

STANDARD DESIGNS

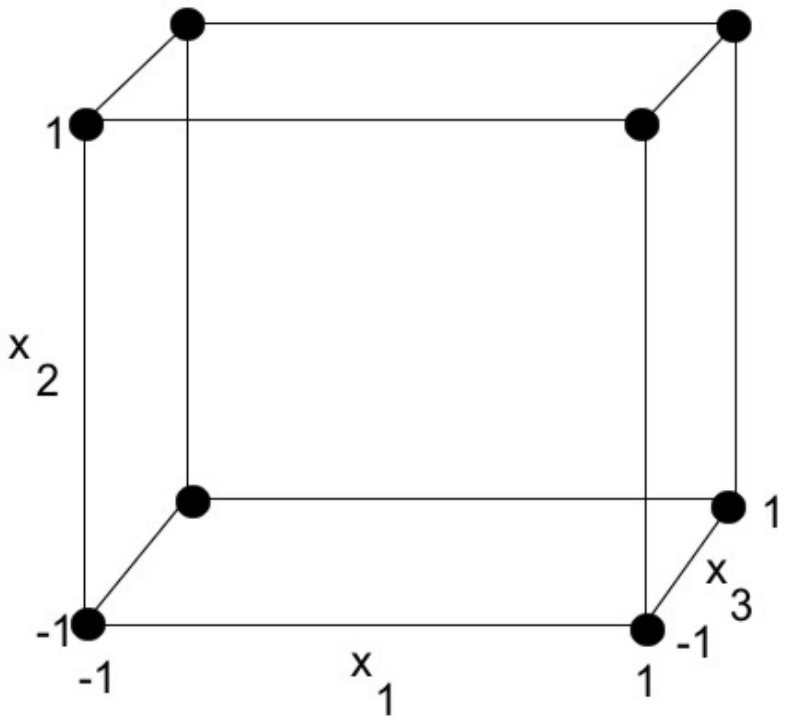
2^3 factorial design

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{23} x_2 x_3 + \beta_{123} x_1 x_2 x_3$$

	x_1	x_2	x_3
1	-1	-1	-1
2	+1	-1	-1
3	-1	+1	-1
4	+1	+1	-1
5	-1	-1	+1
6	+1	-1	+1
7	-1	+1	+1
8	+1	+1	+1

STANDARD DESIGNS

2^3 factorial design cont'd



STANDARD DESIGNS

2-level or 2^{k-q} fractional factorial designs

- Not every combination is tested
- Used for screening purposes
- Full model cannot be estimated
- n is also a power of 2 but we have fewer observations
- q is called the generator of the design

STANDARD DESIGNS

2-level or 2^{k-q} fractional factorial designs Cont'd

- Let's consider a full 2^3 factorial design

Test	I	x_1	x_2	x_3	x_1x_2	x_1x_3	x_2x_3	$x_1x_2x_3$
1	+1	-1	-1	-1	+1	+1	+1	-1
2	+1	+1	-1	-1	-1	-1	+1	+1
3	+1	-1	+1	-1	-1	+1	-1	+1
4	+1	+1	+1	-1	+1	-1	-1	-1
5	+1	-1	-1	+1	+1	-1	-1	+1
6	+1	+1	-1	+1	-1	+1	-1	-1
7	+1	-1	+1	+1	-1	-1	+1	-1
8	+1	+1	+1	+1	+1	+1	+1	+1

STANDARD DESIGNS

2-level or 2^{k-q} fractional factorial designs Cont'd

- We can introduce variable x_4 through interaction $x_1x_2x_3$

Test	I	x_1	x_2	x_3	x_4
1	+1	-1	-1	-1	-1
2	+1	+1	-1	-1	+1
3	+1	-1	+1	-1	+1
4	+1	+1	+1	-1	-1
5	+1	-1	-1	+1	+1
6	+1	+1	-1	+1	-1
7	+1	-1	+1	+1	-1
8	+1	+1	+1	+1	+1

STANDARD DESIGNS

2-level or 2^{k-q} fractional factorial designs Cont'd

...

x_1x_2	x_1x_3	x_1x_4	x_2x_3	x_2x_4	x_3x_4	$x_1x_2x_3$	$x_1x_2x_4$	$x_1x_3x_4$	$x_2x_3x_4$	$x_1x_2x_3x_4$
+1	+1	+1	+1	+1	+1	-1	-1	-1	-1	+1
-1	-1	+1	+1	-1	-1	+1	-1	-1	+1	+1
-1	+1	-1	-1	+1	-1	+1	-1	+1	-1	+1
+1	-1	-1	-1	-1	+1	-1	-1	+1	+1	+1
+1	-1	-1	-1	-1	+1	+1	+1	-1	-1	+1
-1	+1	-1	-1	+1	-1	-1	+1	-1	+1	+1
-1	-1	+1	+1	-1	-1	-1	+1	+1	-1	+1
+1	+1	+1	+1	+1	+1	+1	+1	+1	+1	+1

- We note that the following columns are the same:
  I and $x_1x_2x_3x_4$; x_2 and $x_1x_3x_4$; x_3 and $x_1x_2x_4$; x_1x_2 and x_3x_4 ; x_1x_3 and x_2x_4 ; x_2x_3 and x_1x_4 ; $x_1x_2x_3$ and x_4.

STANDARD DESIGNS

2-level or 2^{k-q} fractional factorial designs Cont'd

Table: Resolution numbers of fractional factorial experiments

Resolution	Main effects aliased with:	2-factor interactions aliased with:
III	2-factor interactions or higher	Main effects and interactions
IV	3-factor interactions or higher	2-factor interactions or higher
V	4-factor interactions or higher	3-factor interactions or higher

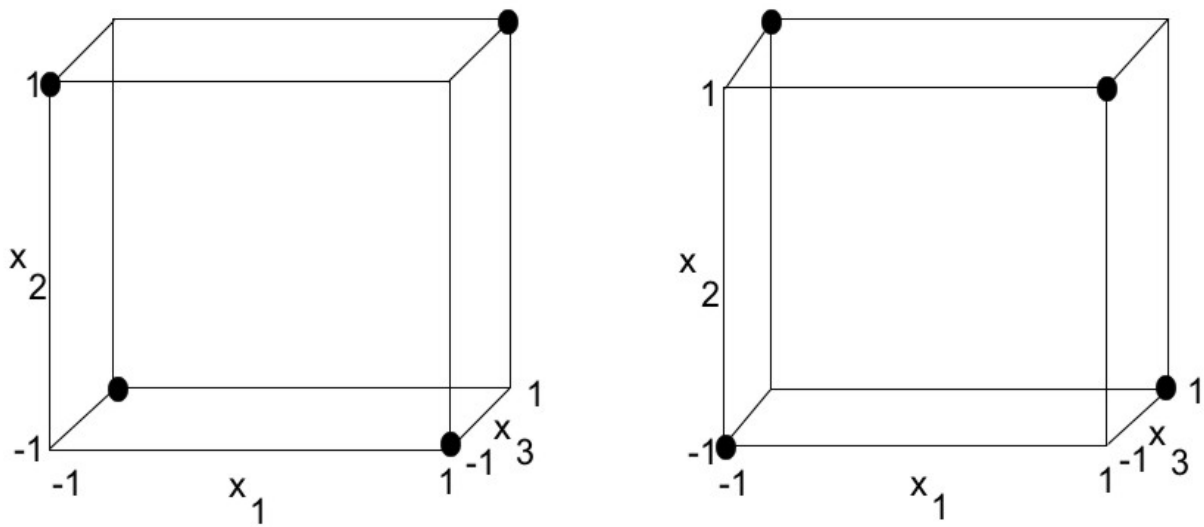
Notation: e.g 2^{3-1}_{III} or 2^{3-1}_{IV}

STANDARD DESIGNS

2-level or 2^{k-q} fractional factorial designs Cont'd

2^{3-1} Fractional Factorial Design

$(+1 = x_1x_2x_3)$ $(-1 = x_1x_2x_3)$



STANDARD DESIGNS

Plackett-Burman Designs

- Only meant for estimating main effect models

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k$$

- n is a multiple of 4
- They are very efficient screening designs

STANDARD DESIGNS

8-point Plackett-Burmann Design for k=7

	x_1	x_2	x_3	x_4	x_5	x_6	x_7
1	+1	+1	+1	-1	+1	-1	-1
2	-1	+1	+1	+1	-1	+1	-1
3	-1	-1	+1	+1	+1	-1	+1
4	+1	-1	-1	+1	+1	+1	-1
5	-1	+1	-1	-1	+1	+1	+1
6	+1	-1	+1	-1	-1	+1	+1
7	+1	+1	-1	+1	-1	-1	+1
8	-1	-1	-1	-1	-1	-1	-1

STANDARD DESIGNS

Two - Level Designs- Conclusion

- all model parameters are estimated independently of each other
- all these designs can be used for quantitative and qualitative variables

STANDARD DESIGNS

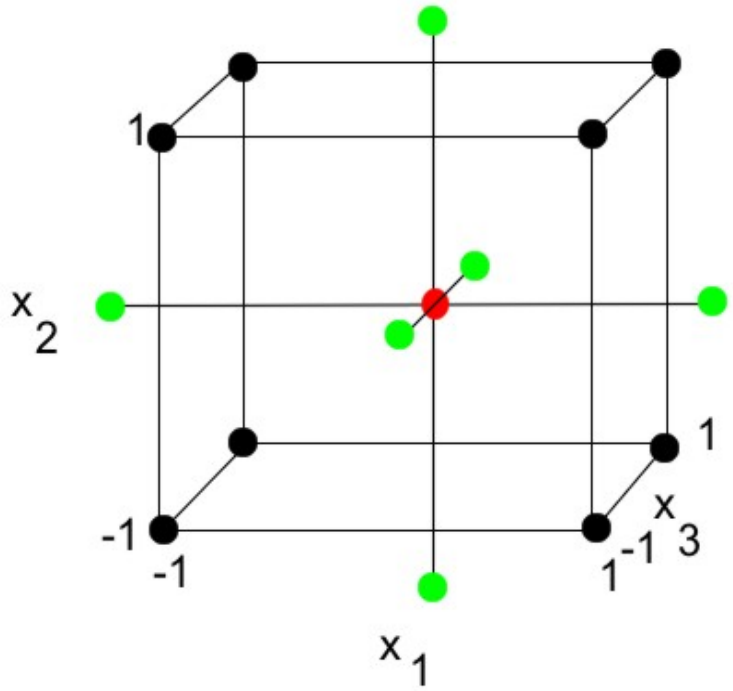
Central Composite Design (CCD)

- Box & Wilson (1951)
- 3 Components:
  - 2^k factorial (or 2^{k-q} fractional factorial) design
  - $2k$ axial or star points (at a distance α from center)
  - at least one center point
    - ★ $\alpha = 1$: face-centered CCD
    - ★ $\alpha = \sqrt{k}$: Spherical CCD

STANDARD DESIGNS

Central Composite Design Cont'd

CCD for $\alpha = 1.68 > 1$



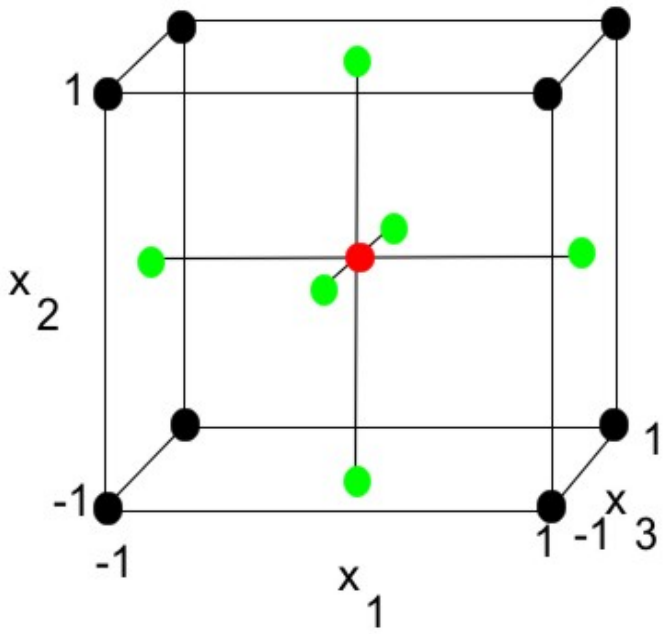
CCD for $\alpha = 1.68 > 1$ cont'd

run	x_1	x_2	x_3
1	-1	-1	-1
2	-1	-1	1
3	-1	1	-1
4	-1	1	1
5	1	-1	-1
6	1	-1	1
7	1	1	-1
8	1	1	1
9	-1.68	0	0
10	1.68	0	0
11	0	-1.68	0
12	0	1.68	0
13	0	0	-1.68
14	0	0	1.68
15	0	0	0
16	0	0	0

STANDARD DESIGNS

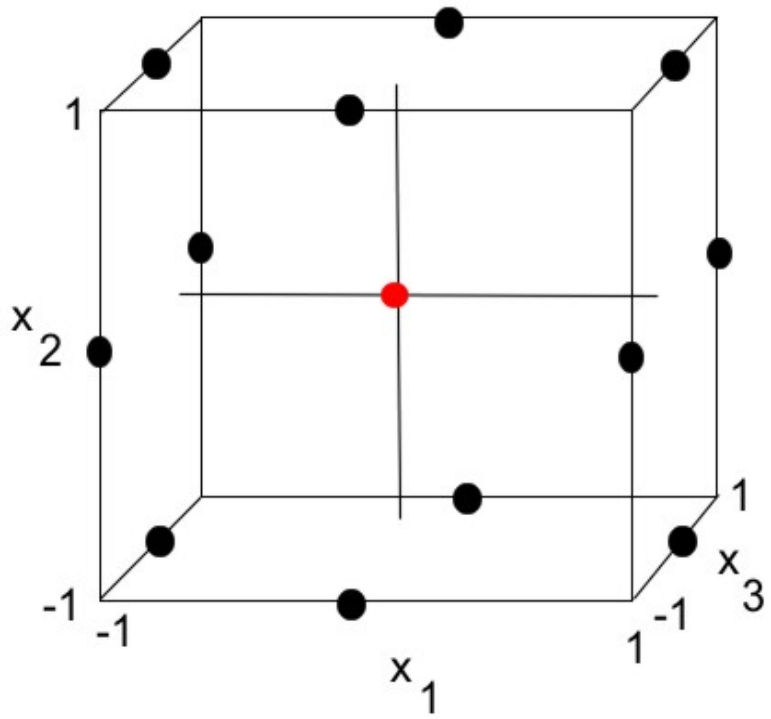
Central Composite Design Cont'd

Face- Centered CCD i.e. $\alpha = 1$



STANDARD DESIGNS

Box- Behnken Design



Box- Behnken Design Cont'd

run	x_1	x_2	x_3
1	-1	-1	0
2	-1	1	0
3	1	-1	0
4	1	1	0
5	0	-1	-1
6	0	-1	1
7	0	1	-1
8	0	1	1
9	-1	0	-1
10	1	0	-1
11	-1	0	1
12	1	0	1
13	0	0	0
14	0	0	0
15	0	0	0

STANDARD DESIGNS

Latin Square designs

A Latin square is an $n \times n$ array filled with n different Latin letters, each occurring exactly once in each row and exactly once in each column.

HYPOTHESIS TESTING

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

A hypothesis is an assumption or a claim about some characteristic of a population, which we should be able to support or reject on the basis of empirical evidence.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

For example, an electric bulb manufacturing company may claim that the average life of its bulbs is at least 1000 hours.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

Hypothesis testing is a process for choosing between different alternatives. The alternatives have to be mutually exclusive and exhaustive.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

Being mutually exclusive means when one is true, the other is false and vice versa. Being exhaustive means that there should not be any possibility of any other relationship between the parameters.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

In the example of the electric bulb manufacturer, the following two options will have to be considered to verify the manufacturer's claim.

1. Average life of the bulb is greater than or equal to 1000 hours.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

2. Average of the bulb is less than or equal to 1000 hours.

We can see that these options are mutually exclusive as well as exhaustive.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

Typically, in hypothesis testing, we have two options to choose from. These are termed as null hypothesis and alternative hypothesis.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

Null Hypothesis (H_o) - It is the presumption that is accepted as correct unless there is strong evidence against it.

Alternative Hypothesis (H_1 or H_A) – It is accepted when H_o is rejected.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

In other words, H_o is the status quo and H_1 is the negation of the status quo.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

Type I and Type II errors

- Type I: If we reject H_o when it should be accepted.
- Type II: If we accept H_o when it should be rejected.

One way to reduce both types of errors is to increase the sample size.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

The probabilities of the two errors are given special symbols

$$\alpha = P(\text{type I error}) = \text{P}(\text{reject } \text{H}_0 \,|\, \text{H}_0 \text{ is true})$$

$$\beta = P(\textit{type II error}) = \text{P}(\text{Accept } \text{H}_0 \,|\, \text{H}_0 \text{ is false})$$

Sometimes it is convenient to work with the power of the test, where

$$\textit{Power} = 1 - \beta = \text{P}(\text{reject } \text{H}_0 \,|\, \text{H}_0 \text{ is false})$$

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

- The general procedure in hypothesis testing is to specify a value of the probability of type I error α, often called the *significance* of the test, and then design the test procedure so that the probability of type II error β has a suitably small value.

Significance level (p-value)

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

There is always a probabilistic component involved in the accept-reject decision in testing hypothesis. The criterion that is used for accepting or rejecting a null hypothesis is called significance level or p-value

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

The p-value represents the probability of concluding (incorrectly) that there is a difference in your samples when no difference exists.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

It is a statistic calculated by comparing the distribution of a given a sample data and an expected distribution (normal, F, t, etc) and it is dependent on the statistical test being performed.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

Note:

P-value of 0.05 means there is only 5% chance that you would be wrong in concluding that the population are different or 95% confident of making a right decision.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

COMPARING MEANS:

One or Two Samples t-Tests

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

Introduction

T-tests are commonly used when making comparisons between the means of two samples or between some standard value and the mean of one sample.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

There are different varieties of t-tests which are used in different conditions depending on the design of the experiment or the nature of the data.

Types of t-tests

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

Three different types:

- One sample t-test,
- Independent samples t-test
- Dependent (paired) samples t-test.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

One Sample t-test

- One sample t-test is used to compare the mean of a single sample with the population mean. Some situations where one sample t-test can be used are as follows

- An economist wants to know if the per capita income of a particular region is same as the national average.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

- A market researcher wants to know if the proposed product will be able to penetrate to a certain level in the households in order to make its introduction profitable.

- The Quality Control department wants to know if the mean dimensions of a particular product have shifted significantly away from the original specifications.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

Independent Samples t-test

- In many real life situations, we cannot determine the exact value of the population mean. We are only interested in comparing two populations using a random sample from each. Such experiments, where we are interested in detecting differences between the means of the two independent groups are called independent samples test.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

- A labor union wants to compare the the productivity levels of workers for two different groups.
- An aspiring MBA student wants compare the salaries offered to the graduates of two business schools.

Dependent (Paired) Samples t-test

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

In the case of independent sample test for testing the differences between means, we assume that the observations on one sample are not dependent on the other. However, this assumption limits the scope of analysis as in many cases the study has to be done on the same set of element to control some of the sample specific extraneous factors.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

Such experiments where the observations are made on the same sample at two different times, is called dependent or paired sample t-test.

Some situations where Dependent Samples t-test can be used are as follows

Created with the trial version of Foxit Advanced PDF Editor
To remove this notice, visit: www.foxitsoftware.com/shopping

- The HR manager wants to know if a particular training program had any impact in increasing the motivation level of the employees
- The Production manager wants to know if a new method of handling machines helps in reducing the break down period.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

- An educationist wants to know if interactive teaching helps students learn more as compared to one-way lecturing.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

Examples:1. One Sample t-test.

- A business school in its advertisements claims the average salary of its graduates in a particular lean year is at par with the average salaries offered at the top five business schools. A sample of 30 graduates, from the business school whose claim was to be verified, was taken at random. The average salary of the top five business schools was given as Rs 750,000.The salaries are given in table 4.1.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

Table 4.1

Graduate Student	Salary (in Rs '000)	Graduate Student	Salary (in Rs '000)
1	750	16	770
2	600	17	680
3	600	18	670
4	650	19	740
5	700	20	760
6	780	21	775
7	860	22	845
8	810	23	870
9	780	24	640
10	670	25	690
11	690	26	715
12	550	27	630
13	610	28	685
14	715	29	780
15	755	30	635

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

2. Independent Samples t-test

- A study was conducted to compare the efficiency of the workers of two mines, one with private ownership and the other with government ownership. The researcher was of the view that there is no significant difference in their efficiency levels.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

Table 4.2

Miner	Mine	Output (in tones)	Miner	Mine	Output (in tones)
1	1	48	23	2	41
2	1	45	24	2	39
3	1	33	25	2	35
4	1	39	26	2	34
5	1	34	27	2	33
6	1	49	28	2	36
7	1	33	29	2	37
8	1	45	30	2	37
9	1	48	31	2	41
10	1	44	32	2	42
11	1	45	33	2	39
12	1	45	34	2	38
13	1	36	35	2	38
14	1	48	36	2	39
15	1	41	37	2	41
16	1	47	38	2	40
17	1	39	39	2	41
18	1	49	40	2	40
19	1	38	41	2	38
20	1	45	42	2	41
21	2	42	43	2	43
22	2	44	44	2	40

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

- In the table we get results of the two tests-Levene's Test for Equality of Variances and the t-test for Equality of Means.

- The table contains two sets of analysis, the first one assuming equal variances in the two groups and the second one assuming unequal variances.

- The Levene's test tells us which statistic to consider to analyze the equality of the means.
- It tests the null hypothesis that the two groups have equal variances.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

- A small value of significance associated with Levene's test indicates that the two groups have unequal variances and the null hypothesis is false.

- In the given example, a very small value of this test statistics indicates that the two groups, *Mine 1* and *Mine 2*, do not have equal variance.

- Therefore, the statistics associated with *equal variances not assumed* should be used for t-test for Equality of Means.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

- The t-test result (with equal variances not assumed) shows t statistic of 2.535 with 27.03 degrees of freedom.

- The corresponding two-tailed p-value is 0.017, which is less than 0.05 but higher than 0.01.

- Therefore, we can reject the null hypothesis at 5% significance level, which means that the average outputs of the two mines are significantly different from each other, i.e.,the miners of the two mines do not have the same efficiency

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

- However at 1% significant level, the null hypothesis will have to be accepted since the p-value is greater than 0.01.

- This means that at 1% significant level, the claim that the efficiency of the miners of the two mines is same is right.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

One-tailed and two-tailed Tests

- We may recall the hypotheses we tested above:

$$H_0 : 01 = 02$$

$$H_1 : 01 \neq 02$$

- Here we are only interested in knowing if the efficiency of the workers of mine 1 is same as that of the workers of mine 2.

- The appropriate test for this would be a two-tailed test. However, if we were to test the assumption that the efficiency of workers of mine 1 is greater than that of the workers of mine 2, the null and alternative hypotheses would be different and we would need the p-value of one-tailed test.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

- The new hypotheses would be:

$$H_0 : 01 \leq 02$$

$$H_1 : 01 > 02$$

- The one-tailed significance value or p-value can be obtained by dividing the two-tailed value by two. Thus the one-tailed p-value in this case would be 0.0085, which is less than 0.01.
- Therefore we reject the null hypothesis even at 1% significance level and conclude that the efficiency of the workers of mine 1 is greater than that of the workers of mine 2.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

3. Dependent Samples t-test

- A corporate training claimed that its training program can greatly enhance the efficiency of call centre employees. A big call center sent some of its employees for the training program. The efficiency was measured by the number of deals closed by each employee in one-month period.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

Data was collected for a one-month period before sending the employers for the training program. After the training program, data was again collected on the same employees for a one-month period. The data is given in Table 4.3.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

Table 4.3

Employee	Before Training Program	After Training Program	Employee	Before Training Program	After Training Program
1	41	44	11	46	39
2	35	36	12	42	40
3	40	48	13	37	36
4	50	47	14	34	39
5	39	40	15	38	50
6	45	52	16	42	46
7	35	35	17	46	49
8	36	51	18	39	42
9	44	46	19	40	51
10	40	55	20	45	37

What If There Are More Than Two Factor Levels?

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

- The *t*-test does not directly apply
- There are lots of practical situations where there are either more than two levels of interest, or there are several factors of simultaneous interest
- The analysis of variance (ANOVA) is the appropriate analysis "engine" for these types of experiments
- The ANOVA was developed by Fisher in the early 1920s, and initially applied to agricultural experiments
- Used extensively today for industrial experiments

COMPLETELY RANDOMIZED DESIGNS (CRDs)

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

- A completely randomized design is a design for which independent random samples of experimental units are selected for each treatment.
- We use CRD to refer to both designed and observational experiments. Thus, the only requirement is that the exptal units to which treatments are applied (designed) or on which treatments are observed (observational) are independently selected for each treatment.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

ASSUMPTIONS

1. Samples are selected randomly and independently from the respective populations.

2. All k population probability distributions are normal.

3. The k population variances are equal.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

- Let us now compare the cleansing action of three detergents on the basis of the whiteness readings made on 15 swatches of white cloth, which were first soiled with India ink and then washed in an agitator-type machine with the respective detergents.
- Detergent A: 77, 81, 71, 76, 80
- Detergent B: 72, 58, 74, 66, 70
- Detergent C: 76, 85, 82, 80, 77
- The means of these three samples are 77, 68, 80 and we want to know whether the differences among them are significant or whether they can be attributed to chance

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

- In general, in a problem like this, we have independent random samples of size n from
- k populations. The j^{th} value from the i^{th} population is denoted by x_{ij}, that is

$$\begin{aligned}
&Population\ 1 \ : \ x_{11},\ x_{12},.....,x_{1n}\\
&Population\ 2 \ : \ x_{21},\ x_{22},.....,x_{2n}\\
&\quad .\qquad\quad .\quad .\qquad\quad .\\
&\quad .\qquad\quad .\quad .\qquad\quad .\\
&\quad .\qquad\quad .\quad .\qquad\quad .\\
&Population\ k \ : \ x_{k1},\ x_{k2},.....,x_{kn}
\end{aligned}$$

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

- We assume that the corresponding random variables, which are all independent, have normal distribution with the respective means and common variance .
- The model for the observations can therefore be given as

$$x_{ij} = \mu_i + \varepsilon_{ij}, \quad i = 1, 2.....k \ \ and \ \ j = 1, 2.....n$$

Where $\varepsilon_{ij} \sim N(0, \sigma^2)$ *are the values of* nk *independent random* variables.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

- To allow the generalization of this model to more complicated kinds of situations, it is usually written as

$$x_{ij} = \mu + \alpha_i + \varepsilon_{ij}, \quad i = 1, 2, \ldots\ldots k \; ; \; j = 1, 2 \ldots . n$$

$$\mu = grand \;\; mean$$

$$\alpha_i = treatment \;\; effects, \;\; such \;\; that \;\; \sum_{i=1}^{k} \alpha_i = 0$$

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

- In practice, why do we refer to the different populations as different treatments?

Many ANOVA techniques were originally developed in connection with agricultural experiments where different fertilizers, for example were regarded as different treatments applied to the soil.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

- 3 detergents are 3 different treatments. In other examples,
- 4 nationalities = 4 different treatments
- 5 advertising campaigns =5 different treatments

- *Note*: “Levels” is another term often used instead of “treatments

HYPOTHESIS

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

- The objective of a CRD is usually to compare the treatment means.
- A CRD is a one-way design. Hence, we use one-way analysis of variance or one-way ANOVA.

If we denote the true, or population, means of the k treatments as $\mu_1, \mu_2, ..., \mu_k$, then we will test the null hypothesis (H_0) that the treatment means are all equal against the alternative hypothesis (H_1) that at least two of the treatment means differ or the population means are not all equal:

$H_0: \mu_1 = \mu_2 = ...\mu_k$, or

$H_0: \alpha_i = 0 \; for \; i = 1, 2, ..., k$

$H_1: \mu_i \neq \mu_j$ for at least one pair (i, j); (i.e at least two of the k treatment means differ) , or

$H_1: \alpha_i \neq 0 \; for \; at \; least \; one \; value \; of \; i$

SUM OF SQUARES

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

- The test itself is based on an analysis of the total variability of the combined data.
- The total variability is called the total sum of squares(SST)

- If the null hypothesis is true, all this variability is due to chance, but if it is not true, then part of the above sum of squares is due to the differences among the population means.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

- The 2 contributors to the total variability are called:
  - Treatment Sum of Squares (SS(Tr))
    - Error Sum of Squares (SSE), where “error” denotes experimental error or chance variability

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

- The SS(Tr) measures chance variation when the null hypothesis is true, but it also reflects the variation among the population means when the null hypothesis is false.

- The SSE component measures chance variation (the variation within the samples).

THEOREM

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

$$SST = SS(Tr) + SSE$$

$$\sum_{i=1}^{k}\sum_{j=1}^{n}(X_{ij}-\bar{X}_{\bullet\bullet})^2 = n\sum_{i=1}^{k}\left(\bar{X}_{i\bullet}-\bar{X}_{\bullet\bullet}\right)^2 + \sum_{i=1}^{k}\sum_{j=1}^{n}\left(X_{ij}-\bar{X}_{i\bullet}\right)^2$$

$$\text{where } \bar{X}_{i\bullet} = \frac{X_{i\bullet}}{n} \text{ is the mean of the observations from the } i^{th} \text{ population}$$

$$X_{i\bullet} = \sum_{j=1}^{n} X_{ij}$$

NOTE: The "dot" subscript notation implies summation over the subscript that it replaces.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

$$\bar{X}_{\bullet\bullet} = \frac{X_{\bullet\bullet}}{kn} \text{ is the sum of all } k \cdot n \text{ observations}$$

$$X_{\bullet\bullet} = \sum_{i=1}^{k} \sum_{j=1}^{n} X_{ij}$$

$$\bar{X}_{\bullet j} = \frac{X_{\bullet j}}{k}$$

$$X_{\bullet j} = \sum_{i=1}^{k} X_{ij}$$

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

SIMPLIFIED NOTATION

$$SS(Tr) = \sum_{i=1}^{k} (\bar{x}_i - \bar{x})^2$$

where $\bar{x}$ = Overall mean of all sample measurements

$$SSE = \sum_{j=1}^{n_1} (x_{1j} - \bar{x}_1)^2 + \sum_{j=1}^{n_2} (x_{2j} - \bar{x}_2)^2 + \ldots + \sum_{j=1}^{n_k} (x_{kj} - \bar{x}_k)^2$$

This formula can be simplified by recalling the formula for the sample variance, s^2.

$$S^2 = \sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{n-1}$$

Note that each sum in SSE is simply the numerator of s^2 for that particular treatment.

Consequently, we can rewrite SSE as

$$\text{SSE} = (n_1 - 1)s_1^2 + (n_2 - 1)s_2^2 + \ldots\ldots + (n_k - 1)s_k^2$$

where $s_1^2, s_2^2, \ldots\ldots, s_k^2$ are the sample variances for the k treatments.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

To make the two measurements of variability comparable, we divide each by the the degrees of freedom to convert the sums of squares to mean squares.

Mean Square for treatments, MS (Tr), which measures the variability among the treatment means, is given as

$$MS(Tr) = \frac{SS\ (Tr)}{k-1}.$$

Mean Square for Error, MSE, which measures the sampling variability within the treatments is given as

$$MSE = \frac{SSE}{k\ (n-1)}$$

F–VALUE

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

$$F = \frac{MS(Tr)}{MSE} \text{ is a value of a random variable having}$$

F – distribution with k -1 and $k(n$ -1) degrees of freedom.

We reject H_0 that the population means are all equa if the value we obtain from F exceeds $F_{\alpha,k\text{-}1,k(n\text{-}1)}$

where α is the level of significance

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square	F
Treatments	$k-1$	$SS(Tr)$	$MS(Tr)$	$\frac{MS(Tr)}{MSE}$
Error	$k(n-1)$	SSE	MSE	
Total	$kn-1$	SST		

- NOTE:

  n= the number of samples in each population

EXAMPLE

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

Suppose the USGA wants to compare the mean distances associated with four different brands of golf balls when struck with a driver.

A completely randomized design is employed, with Iron Byron, the USGA's robotic golfer, using a driver to hit a random sample of 10 balls of each brand in a random sequence.

The distance is recorded for each hit, and the results are shown in the following table.

Results of Completely Randomized Design : Iron Byon Driver

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

	BRAND A	BRAND B	BRAND C	BRAND D
	251.2	263.2	269.7	251.6
	245.1	262.9	263.2	248.6
	248.0	265.0	277.5	249.4
	251.1	254.5	267.4	242.0
	260.5	264.3	270.5	246.5
	250.0	257.0	265.5	251.3
	253.9	262.8	270.7	261.8
	244.6	264.4	272.9	249.0
	254.6	260.6	275.6	247.1
	248.8	255.9	266.5	245.9
Sample Mean	250.78	261.06	269.95	249.32

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

1. Set up the test to compare the mean distances for the four brands. Use $\alpha = 0.10$

2. Use the SPSS program to obtain the test statistic and p-value. Interpret the results.

Solution

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

a) To compare the mean distances of the four brands, we first specify the hypotheses to be tested. Denoting the population mean of the i^{th} brand by μ_i . Test

$$H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$$

H_1 : The mean distances differ for at least two of the brands

Test Statistic, F= $\dfrac{MS(Tr)}{MSE}$

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

From the table

$k = 4,\ nk = 40,\ \bar{x} = 257.78$

$$SS(Tr) = \sum_{i=1}^{k} n_i (\bar{x}_i - \bar{x})^2$$

$$= 10(250.78 - 257.78)^2 + 10(261.06 - 257.78)^2 + 10(269.95 - 257.78)^2 + 10(249.32 - 257.78)^2$$

$$= 2,794.39$$

$$MS(Tr) = \frac{SS(Tr)}{k-1}$$

$$= \frac{2,794.39}{4-1}$$

$$= 931.46$$

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

$$SSE = \sum_{j=1}^{n_1}(x_{1j} - \bar{x}_1)^2 + \sum_{j=1}^{n_2}(x_{2j} - \bar{x}_2)^2 + \ldots\ldots\ldots + \sum_{j=1}^{n_k}(x_{kj} - \bar{x}_k)^2$$

$$= (251.2\text{-}250.78)^2 + \ldots\ldots\ldots + (248.8 - 250.78)^2$$

$$+ (263.2\text{-}261.06)^2 + \ldots\ldots\ldots + (255.9 - 261.06)^2$$

$$+ (269.7\text{-}269.95)^2 + \ldots\ldots\ldots + (266.5 - 269.95)^2$$

$$+ (251.6\text{-}249.32)^2 + \ldots\ldots\ldots + (245.9 - 249.32)^2$$

$$= 762.30$$

ALTERNATIVELY,

$$\text{SSE} = (\text{n}_1 - 1)s_1^2 + (n_2 - 1)s_2^2 + (n_3 - 1)s_3^2 + (n_4 - 1)s_4^2$$

$$= 9(4.735163965)^2 + 9(3.866149391)^2 + 9(4.500925831)^2$$

$$+ 9(5.203161432)^2$$

$$= 762.3009$$

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

$$MSE = \frac{SSE}{k(n-1)}$$

$$= \frac{762.30}{4(10-1)}$$

$$= 21.18$$

$$F = \frac{MS(Tr)}{MSE} = 43.99$$

$$F_{\alpha,k-1,k(n-1)} = F_{0.10,3,36} = 2.25$$

$$F > F_{0.10,3,36}$$

Therefore we reject H_0 and conclude that at least two of the brands differ with respect to mean distance travelled when struck by the driver.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

Example Using SPSS

An oil company has introduced a new brand of gasoline in its outlets in three major metro cities. However, they are not sure how the new brand is selling at the three places since there is a lot of difference in the driving habits of people in the three metros. The company selected 10 outlets in each city and tabulated the data on an average daily sale at each of the selected outlets. The data is presented in Table 5.1. The items 1, 2 and 3 in the table represent the three metros.

Table 5.1

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

Outlets	Metro	Volume (in gallons)	Outlets	Metro	Volume (in gallons)	Outlets	Metro	Volume (in gallons)
1	1	15.00	11	2	8.00	21	3	14.60
2	1	13.50	12	2	8.25	22	3	16.50
3	1	14.20	13	2	6.40	23	3	13.20
4	1	13.00	14	2	7.20	24	3	13.60
5	1	15.00	15	2	4.40	25	3	21.00
6	1	11.00	16	2	11.30	26	3	11.70
7	1	12.50	17	2	7.80	27	3	12.30
8	1	16.00	18	2	8.10	28	3	17.00
9	1	14.25	19	2	9.00	29	3	14.75
10	1	13.50	20	2	6.80	30	3	18.00

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

The null hypothesis in this case is:

H_{01} : The average sale of the new brand of the gasoline is same in all the metro cities.

We will also explain multiple comparisons in this example by testing for the following null hypothesis .

H_{02} : The average sale of the new brand of the gasoline in city 1 is same as that of city 2.

H_{03} : The average sale of the new brand of the gasoline in city 1 is same as that of city 2 and city 3 together.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

Please note the difference between H_{02} ,H_{03} and H_{01}. The F-statistics obtained form ANOVA only tells us whether there is any significant difference in the mean values of the three groups, which is our H_{01}. From this, we cannot conclude anything about specific hypotheses as stated in H_{02} and H_{03} .

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

POST-HOC MULTIPLE COMPARISON

Rejection of the null hypothesis in ANOVA only tells us that all population means are not equal. Multiple comparisons are used to assess which group means differ from which others, once the overall F-test shows that at least one difference exists. Many tests are listed under "Post-Hoc" in SPSS. Tukey HSD (honestly significant difference) test is one of the most conservative and commonly-used test.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

Contrast Analysis

Contrasts are also used to do multiple comparisons in ANOVA.

Using contrasts, we may test the significance of specified differences in particular parts of the design, like testing the differences between two particular means or comparing the average of a combination of means to a particular mean.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

Unplanned and Planned Comparisons

The F-statistic in ANOVA only tells us that the dependent variable varies for different levels of factor(s). The case of more than two levels, F-statistic does not tell us the exact way in which the dependent variable differs by the levels of factor(s).

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

To gain a deeper understanding of the subject, we have to rely on unplanned and planned comparisons.

Unplanned Comparison

Unplanned comparison is exploring the differences in means between all possible pairs of groups.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

In the example given above, an unplanned comparison would mean that the researcher is interested in identifying whether the average of gasoline in a particular city is different from any other city.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

While making such comparisons, the researcher has no *a priori* hypotheses about differences in the dependent variable for different levels of the factor(s). Post-Hoc tests are used for such unplanned comparison.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

Planned comparisons, on the other hand, are decided *a priori* by the researcher, based on theoretical considerations.

Such comparisons can be done using the Contrasts analysis.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

A common and most frequent-used technique is to use a linear combination of contrasts. The method of assigning weights is very simple and can be learned easily with little practice.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

We had three null hypotheses. We tested the first of these (H_{01}) using the F-statistic as illustrated earlier. The other two (H_{02} and H_{03}) are based on planned comparisons.

For H_{02} ,the contrast coefficients will be -1,1 and 0 for city 1,2 and 3 respectively.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

The assignment compares the mean for city 1 with the mean for city 2.

Likewise, for H_{03} , the contrasts coefficients will be 2,-1 and -1 for city 1, 2 and 3 respectively.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

This assignment compares the mean for city 1with mean for city 2 and 3 together.

Essentially, if there are three factors (A,B,C), we can make the following comparisons:

1. We can compare one factor with another factor (A with B, B with C or A with C). The corresponding contrasts will be (1,-1,0), (0,1,-1) and (1, 0,-1).

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

2. We can compare one factor with two other factors (A with B and C……). The corresponding contrasts will be (2,-1,-1),or (-2,1,1).

The assignment of contrasts will be similar if there are four factors. In the case of four factors, if we have to compare the mean of the first and third, with the mean of the second and fourth, the contrasts will be (1,-1,1,-1).

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

Likewise , if we have to compare the mean of the third factor with the other three, the contrast will be (-1,-1,3,-1).

One has to ensure that the comparisons are orthogonal, while performing more than one planned comparison on the same set of data. This condition is satisfied if the products of coefficients assigned to each level sum to zero.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

RANDOMIZED-BLOCK DESIGNS

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

If the completely randomized design results in nonrejection of the null hypothesis that the treatment means differ because the sampling variability (as measured by MSE) is large, we may want to consider an experimental design that better controls the variability.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

In contrast to the selection of independent samples of experimental units specified by the completely randomized design, RBD utilizes experimental units that are *matched sets,* assigning one from each set to each treatment.

The matched sets of experimental units are called *blocks*.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

The theory behind the randomized block design is that the sampling variability of the experimental units in each block will be reduced, in turn reducing the measure of error, MSE.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

Definition:

The **randomized block design** consists of a two-step procedure:

1. Matched sets of experimental units, called **blocks,** are formed, each block consisting of *k* experimental units (where k is the number of treatments). The *b* blocks should consist of experimental units that are as similar as possible.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

2. One experimental unit from each block is randomly assigned to each treatment, resulting in a total of $n = bk$ responses.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

- To do this in the example dealing with the three detergents *(discussed earlier),* we might soil the swatches with exactly equal amounts of India ink, always use the same washing time and water of exactly the same hardness and temperature, and inspect (and, if necessary, adjust) the measuring instruments after each use.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

- Under such rigidly controlled conditions, significant differences among the sample means cannot be due to differently soiled swatches or differences in washing time, water temperature, water hardness, or measuring instruments.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

- The differences among the means only show that the detergents are not all equally effective if they are used in this narrowly restricted way.

- We cannot say whether the same difference will exist if the washing were longer or shorter, if the water had a different temperature or hardness, and so on.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

- To present the theory of the analysis of variance associated with a randomized block design, we shall use terminology similar to that introduced in the CRB, now referring to the two variables as treatments and **blocks**.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

If X_{ij} for $i = 1, 2, \ldots\ldots, k$ and $j = 1, 2, \ldots., b$ are the values of independent random variables having normal distributions with the respective means μ_{ij} and the common variance σ^2, we shall consider the array.

	Block 1	Block 2	……..	Block b
Treatment 1	x_{11}	x_{12}	……..	x_{1b}
Treatment 2	x_{21}	x_{21}	……..	x_{2b}
……..	……	……	……	……
Treatment k	x_{k1}	x_{k1}	……..	x_{kb}

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

- The model for a randomized-block design as

$$x_{ij} = \mu + \alpha_i + \beta_j + e_{ij}$$

$$for\ i = 1, 2, ..., k\ and\ j = 1, 2, ...b.$$

Here μ *is the GRAND MEAN. The treatment effects* α_i *are such that* $\sum_{i=1}^{k} \alpha_i = 0$, *the block effect* β_j *are such that* $\sum_{j=1}^{b} \beta_j = 0$, *and the* e_{ij} *are the values of independent random variables having normal distribution with zero means and the common variance* σ^2.

Note that $\mu_{ij} = \mu + \alpha_i + \beta_j$

$$\frac{\sum_{i=1}^{k}\sum_{j=1}^{b} \mu_{ij}}{kb} = \mu$$

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

The two null hypotheses that we shall want to test are that the treatment effects are equall to zero and the block effects are all equal to zero;

that is

$H_0 : \alpha_i = 0$ for i=1,2,.....k

and

$H_0^{'} : \beta_j = 0$ for j=1,2,.....b

The alternate to H_0 is that the treatment effects are not all equal to zero, and the alternative to $H_0^{'}$ is that the block effects are not all equal to zero.

Symbolically,

$H_1 : \alpha_i \neq 0$ for at least one value of i

$H_1^{'} : \beta_j \neq 0$ for at least one value of j

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

The two-way analysis is based on the following generalization

$$SST = SS(Tr) + SSB + SSE$$

$$\sum_{i=1}^{k}\sum_{j=1}^{b}(\bar{x}_{ij} - \bar{x}_{..})^2 = b\sum_{i=1}^{k}(\bar{x}_{i.} - \bar{x})^2 + k\sum_{j=1}^{b}(\bar{x}_{.j} - \bar{x}_{..})^2 + \sum_{i=1}^{k}\sum_{j=1}^{b}(x_{ij} - \bar{x}_{i.} - \bar{x}_{.j} + \bar{x})^2$$

where $\bar{x}_{i.}$ *is the mean of the observations for the* i^{th} *treatment,* $\bar{x}_{.j}$ *is the mean of the observations for the* j^{th} *block, and* $\bar{x}_{..}$ *is the mean of all kb observations.*

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

We reject H_0 if $F_{Tr} \geq F_{\alpha,k-1,(b-1)(k-1)}$ where

$$F_{Tr} = \frac{MS(Tr)}{MSE}$$

Re *ject* $H_0^{'}$ *if* $F_B \geq F_{\alpha,b-1,(b-1)(k-1)}$ *where*

$$F_B = \frac{MSB}{MSE}$$

This kind of analysis is called a TWO-WAY ANALYSIS OF VARIANCE

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

SIMPLIFIED NOTATIONS

- $\text{SSB} = \sum_{i=1}^{b} k(\bar{x}_{B_i} - \bar{x})^2$

where $\bar{x}_{B_i}$ represents the sample mean for the ith block and k (the number of treatments) is the number of measurements in each block.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

- $SS(Tr)=\sum_{i=1}^{k} b(\bar{x}_{T_i} - \bar{x})^2$

where $\bar{x}_{T_i}$ represents the sample mean for the ith treatment, b (the number of blocks) is the number of measurements for each treatment, and k is the number of treatments.

- $SST=\sum_{i=1}^{n} (x_i - \bar{x})^2$

where n=bk

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

- SSE= SST – SS(Tr) – SSB

- MS(Tr)= $\frac{SS(Tr)}{k-1}$

- MSE = $\frac{SSE}{(b-1)(k-1)}$

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

- Finally the formula for the test statistic that compares the variation among the (number) treatment means to the sampling variability within each of the treatments is:

$$F_{Tr} = \frac{MS(Tr)}{MSE}$$

- However, we can test the hypothesis that the block means are significantly different, using:

$$F_{B} = \frac{MSB}{MSE}$$

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

Source of Variation	*Degrees of Freedom*	*Sum of Squares*	*Mean Square*	F
Treatments	(k–1)	SS(Tr)	MS(Tr)	$F_{Tr} = \frac{MS(Tr)}{MSE}$
Blocks	(b–1)	SSB	MSB	$F_B = \frac{MSB}{MSE}$
Error	(b–1)(k–1)	SSE	MSE	
Total	kb–1	SST		

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

Example 2

Suppose USGA *(discussed earlier)* wants to compare the mean distances associated with the four brands of golf balls when struck by a driver, but wishes to employ human golfers rather than the robot, Iron Byron. Assume that 10 balls of each brand are to be utilized in the experiment.

a. Explain how a completely randomized design could be employed.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

b. Explain how a randomized block design could be employed.

c. Which design is likely to provide more information about the differences among the brand mean distances?

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

Solution:

a. Since the completely randomized design calls for independent samples, we can employ such a design by randomly selecting 40 golfers and then randomly assigning 10 golfers to each of the four brands.

Finally, each golfer will strike the ball of the assigned brand, and the distance will be recorded

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

b.The randomized block design employs of relatively homogeneous experimental units.

For example, we could randomly select 10 golfers, and permit each golfer to hit four balls, one of each brand, in a random sequence.

Then each golfer is a block, with each treatment (brand) assigned to each block (golfer).

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

c. Because we expect much more variability among distances generated by "real" golfers than by Iron Byron, we would expect the randomized block design to control the variability better than the completely randomized design.

That is, with 40 different golfers, we would expect the sampling variability among the measured distances within each brand to be greater than that among the four distances generated by each of 10 golfers hitting one ball of each brand.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

Now, suppose the randomized block design of part *b* is employed, utilizing a random sample of 10 golfers, with each golfer using a driver to hit four balls, one of each brand, in a random sequence.

a. Set up a test of the research hypothesis that the brand mean distances differ. (The data for the experiment are given in Table below). Use α=0.5

b. Use statistical software to analyse the data, and conduct the test set up in part a.

Golfer (Block)	Brand A	Brand B	Brand C	Brand D
1	202.4	203.2	223.7	203.6
2	242.0	248.7	259.8	240.7
3	220.4	227.3	240.0	207.4
4	230.0	243.1	247.7	226.9
5	191.6	211.4	218.7	200.1
6	247.7	253.0	268.1	244.0
7	214.8	214.8	233.9	195.8
8	245.4	243.6	257.8	227.9
9	224.0	231.5	238.2	215.7
10	252.2	255.2	265.4	245.2
Sample Means	227.0	233.2	245.3	220.0

Edited with the trial version of Foxit Advanced PDF Editor
To remove this notice, visit: www.foxitsoftware.com/shopping

Solution

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

a) To compare the mean distances of the four brands, we first specify the hypotheses to be tested. Denoting the population mean of the i^{th} brand by μ_i . Test

$$H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$$

H_1 : The mean distances differ for at least two of the brands

Test Statistic, F= $\dfrac{MS(Tr)}{MSE}$

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

From the table

$k = 4,\ nk = 40,\ \bar{x} = 231.5725$

$$SS(Tr) = \sum_{i=1}^{k} b(\bar{x}_i - \bar{x})^2$$

$$= 10(227.0 - 231.57255)^2 + 10(233.2 - 231.57255)^2 + 10(245.3 - 231.57255)^2$$
$$+ 10(220.7 - 231.57255)^2$$
$$= 3302.12025$$

$$MS(Tr) = \frac{SS(Tr)}{k-1}$$
$$= \frac{3302.12025}{4-1}$$
$$= 1100.70675$$

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

$$SST = \sum_{i=1}^{n} (X_i - \bar{X})^2$$

$$= (202.4 - 231.5725)^2 + \ldots\ldots\ldots\ldots + (245.2 - 231.5725)^2$$

$$= 15919.2$$

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

$$SSB = 4(208.225 - 231.57255)^2 + 4(247.8 - 231.57255)^2$$

$$+4(223.775 - 231.57255)^2 + 4(236.965 - 231.57255)^2$$

$$+4(205.45 - 231.57255)^2 + 4(253.2 - 231.57255)^2$$

$$+4(214.825 - 231.57255)^2 + 4(243.675 - 231.57255)^2$$

$$+4(227.35 - 231.57255)^2 + 4(254.5 - 231.57255)^2$$

$$= 12075.60145$$

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

SSE=SST - SS(TR) - SSB

SSE=15919.2 - 3298.7 - 12073.9

=546.6

$$MSE = \frac{SSE}{(b-1)(k-1)} = \frac{546.6}{(10-1)(4-1)} = 20.2444$$

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

$$\text{MSB} = \frac{SSB}{b-1}$$

$$= \frac{12075.60145}{9}$$

$$= 1341.7335$$

$$F_b = \frac{MSB}{MSE}$$

$$= \frac{1341.7335}{20.2444}$$

$$= 66.277$$

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

$$F_t = \frac{MST}{MSE}$$

$$= \frac{1100.70676}{20.244}$$

$$= 54.3709$$

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

Source	DF	SS	MS	F
Treatment	3	3302.120	1100.707	54.370
Block	9	12075.602	1341.734	66.277
Error	27	546.600	20.244	
Total	39	15919.200		

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

Example

An MBA aspirant was interested in knowing the impart of educational background(art/commence and science /engineering) on the final placement salaries. He is also aware that the previous work experience has an impart on the salaries.

Therefore, he chooses educational background and work experience has an impact on the salaries. Therefore, he chooses educational background and work experience as two independent variables. Based on educational background, respondents are categorized into two groups- one belonging to the arts stream.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

Based on previous work experience, respondents are again categorized into two groups- one with experience and the other without experience. Since we have two levels of each of the factors, this is called a 2*2 design.

A sample of 30 students is randomly chosen and their salaries from campus recruitment as well as information on educational background, I represents arts and commence and 2 represents the science and engineering stream. For work experience, 1 represents students with experience and 2 represents students without experience. Salary is given in the multiple GH100,000

Students	Eduback	Workexp	Salary	Students	Eduback	Workexp	Salary
1	1	1	8.50	16	2	1	9.80
2	1	1	10.80	17	2	1	10.20
3	1	1	9.70	18	2	1	11.00
4	1	1	8.80	19	2	2	7.80
5	1	1	7.80	20	2	2	7.30
6	1	1	7.50	21	2	1	6.90
7	1	1	7.80	22	2	1	6.10
8	1	1	6.90	23	2	1	6.25
9	1	2	4.50	24	2	2	3.80
10	1	2	4.10	25	2	2	3.20
11	1	2	7.70	26	2	2	5.10
12	1	2	5.50	27	2	2	4.90
13	1	2	5.60	28	2	2	4.65
14	1	2	5.20	29	2	2	4.80
15	1	2	4.10	30	2	1	5.24

Created with Free Screen Capture
Edit Assembly for Video
To remove this notice, visit:
www.fooltsoftware.com/shopping

Exercise

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

A medical device manufacturer produces vascular grafts (artificial veins). These grafts are produced by extruding billets of polytetrafluoroethylene (PTFE) resin combined with a lubricant into tubes. Frequently, some of the tubes in a production run contain small, hard protrusions on the external surface. These defects are known as “flicks.” The defect is cause for rejection of the unit.

The product developer responsible for the vascular grafts suspects that the extrusion pressure affects the occurrence of flicks and therefore intends to conduct an experiment to investigate this hypothesis.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

However, the resin is manufactured by an external supplier and is delivered to the medical device manufacturer in batches.

The engineer also suspects that there may be significant batch-to-batch variation, because while the material should be consistent with respect to parameters such as molecular weight, mean particle size, retention and peck height ratio, it probably isn't due to manufacturing variation at the resin supplier and natural variation in the material.

Therefore , the production developer decides to investigate the effect of four different levels of extrusion pressure on flicks using a randomized complete block design considering batches of resin as blocks.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

The RCBD is shown in the table on the next slide. Note that there are four levels of extrusion pressure (treatments) and six batches of resin (blocks).

Remember that the order in which the extrusion pressures are tested within each block is random. The response variable is yield, or the percentage of tubes in the production run that did not contain any flicks.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

Randomized complete Block Design for the Vascular Graft Experiment

Extrusion Pressure (PSI)	Batch of Resin (Block) 1	2	3	4	5	6	Treatment Total
8500	90.3	89.2	98.2	93.9	87.4	97.9	556.9
8700	92.5	89.5	90.6	94.7	87.0	95.8	550.1
8900	85.5	90.8	89.6	86.2	88.0	93.4	533.5
9100	82.5	89.5	85.6	87.4	78.9	90.7	514.6
Block Totals	350.8	359.0	364.0	362.2	341.3	377.8	$y_{\bullet\bullet} = 2155.1$

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

The Latin Square Design

- These designs are used to simultaneously control (or eliminate) **two sources of nuisance variability.**

  i.e. It systematically allows blocking in two directions. Thus, the rows and columns represent **two restrictions on randomization.**
- A significant assumption is that the three factors (treatments, nuisance factors) **do not interact**
- If this assumption is violated, the Latin square design will not produce valid results
- A Latin square is an n x n array filled with n different Latin letters, each occurring exactly once in each row and exactly once in each column.
- Latin squares are not used as much as the RCBD in industrial experimentation

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

- We introduced the RCBD as a design to reduce the residual error in an experiment by removing variability due to a known and controllable nuisance variable.
- There are several other types of designs (e.g. latin squares) that utilize the blocking principle.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

Examples

1×1

a

2×2

a	b
b	a

3×3

a	b	c
b	c	a
c	a	b

4×4

a	b	c	d
b	c	d	a
c	d	a	b
d	a	b	c

Latin squares are closely related to a popular puzzle called sudoku puzzle that originated in Japan

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

Completion Problems

- The ugly (?)
  a. k. a. sudoku

9					7		1	
	7	4			2		6	5
	1		8	9		4		
2				8		5	9	
1			2		3			6
	4	5		1				2
		7		4	1		2	
3	2		9			8	5	
	6		5					7

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

Completion Problems

- The ugly (?)
  a. k. a. sudoku

9	5	3	4	6	7	2	1	8
8	7	4	1	3	2	9	6	5
6	1	2	8	9	5	4	7	3
2	3	6	7	8	4	5	9	1
1	9	8	2	5	3	7	4	6
7	4	5	6	1	9	3	8	2
5	8	7	3	4	1	6	2	9
3	2	1	9	7	6	8	5	4
4	6	9	5	2	8	1	3	7

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

Example

Among the nine persons interviewed in a poll, three are Northerners, three are Southerners, and three are Westerners. By profession, three of them are teachers, three are lawyers, and three are doctors, and no two of the same profession come from the same part of Ghana. Also, three are NPP members, three are NDC members, and three are PNC members,

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

and no two of the same political affiliation are of the same profession or come from the same part of Ghana. If one of the teachers is a Northerner and a PNC member, another teacher is a Southerner and an NDC member, and one of the lawyers is a Southerner and an NPP member, what is the political affiliation of a doctor who is a Westerner?.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

solution

- This is a typical 3×3 Latin square problem

	T	L	D
N	PNC	NDC	NPP
S	NDC	NPP	PNC
W	NPP	PNC	NDC

- Hence, the political affiliation of a doctor who is a Westerner is NDC

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

Statistical Analysis of the Latin Square Design

- The statistical (effects) model is

$$y_{ijk} = \mu + \alpha_i + \tau_j + \beta_k + \varepsilon_{ijk} \begin{cases} i = 1, 2, \ldots, p \\ j = 1, 2, \ldots, p \\ k = 1, 2, \ldots, p \end{cases}$$

- y_{ijk} is the observation in the *i*th row and *k*th column for the *j*th treatment, μ is the overall mean, α_i is the *i*th row effect, τ_j is the *j*th treatment effect, β_k is the *k*th column effect and ε_{ijk} is the random error.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

- This model is completely additive, that is there is no interraction between rows, columns and treatments

- Only two of three subscripts are needed to denote a particular observation.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

- Sum of squares:

$$SS_T = SS_{Rows} + SS_{Columns} + SS_{Treatments} + SS_E$$

- The degrees of freedom:

$$p^2 - 1 = p - 1 + p - 1 + p - 1 + (p - 2)(p - 1)$$

- The appropriate statistic for testing for no differences in treatment means is

$$F = \frac{MS_{Treatments}}{MS_E} \sim F_{p-1,(p-2)(p-1)}$$

- The statistical analysis (ANOVA) is much like the analysis for the RCBD.
- The analysis for the rocket propellant example follows

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

Analysis of variance for the Latin Square Design

Sorces of variation	Sum of Squares	Degree of Freedom	Mean Square	F
Treatment	$\frac{1}{p}\sum_{j=1}^{p} y_{\bullet j\bullet}^2 - \frac{y_{\bullet\bullet}^2}{N}$	P-1	$\frac{SS_{Treatment}}{p-1}$	
Rows	$= \frac{1}{p}\sum_{i=1}^{p} y_{i\bullet\bullet}^2 - \frac{y_{\bullet\bullet\bullet}^2}{N}$	P-1	$\frac{SS_{Rows}}{p-1}$	
Columns	$\frac{1}{p}\sum_{k=1}^{p} y_{\bullet\bullet k}^2 - \frac{y_{\bullet\bullet\bullet}^2}{N}$	P-1	$\frac{SS_{columns}}{p-1}$	$F = \frac{MS_{Treatment}}{MS_E}$
Error	SS_E(by subtraction)	(p-2)(p-1)	$\frac{SS_E}{(p-2)(p-1)}$	
Total	$\sum_i \sum_j \sum_k y_{ijk}^2 - \frac{y_{\bullet\bullet\bullet}^2}{N}$	$p^2 - 1$		

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

EXAMPLE

Suppose that an experimenter is studying the effects of five different formulations of a rocket propellant used in aircrew escape systems on the observed burning rate.

Each formulation is mixed from a batch of raw material that is only large enough for five formulations to be tested. Furthermore, the formulations are prepared by several operators, and there may be substantial differences in the skills and experience of the operators.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

Thus, it would seem that there are two nuisance factors to be "averaged out" in the design: the batches of raw material and the operators. The appropriate design for this problem consists of testing each formulation exactly once in each batch of raw material and for each formulation to be prepared exactly once by each of five operators. The five formulations are denoted by A,B,C,D and E

The Rocket Propellant Problem – A Latin Square Design

Edited with the trial version of Foxit Advanced PDF Editor
To remove this notice, visit: www.foxitsoftware.com/shopping

Batches of Raw Material	Operators 1	2	3	4	5
1	A= 24	B= 20	C= 19	D= 24	E= 24
2	B= 17	C= 24	D= 30	E= 27	A= 36
3	C= 18	D= 38	E= 26	A= 27	B= 21
4	D= 26	E= 31	A= 26	B= 23	C= 22
5	E= 22	A= 30	B= 20	C= 29	D= 31

This is a 5x5 Latin square design.

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

We Code the data in the Rocket Propellant Problem by subtracting 25 from each observation

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

Coded Data for the Rocket Propellant Problem

Batches of Raw Material	Operators					$y_{i\bullet\bullet}$
	1	2	3	4	5	
1	A= -1	B= -5	C= -6	D= -1	E= -1	-14
2	B= -8	C= -1	D= 5	E= 2	A= 11	9
3	C= -7	D= 13	E= 1	A= 2	B= -4	5
4	D= 1	E= 6	A= 1	B= -2	C= -3	3
5	E= -3	A= 5	B= -5	C= 4	D= 6	7
$y_{\bullet\bullet k}$	-18	18	-4	5	9	$y_{\bullet\bullet\bullet}$ =10

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

$$SS_T = \sum_i \sum_j \sum_k y_{ijk}^2 - \frac{y_{\bullet\bullet\bullet}^2}{N}$$

$$= 680 - \frac{(10)^2}{25} = 676$$

$$SS_{Batches} = \frac{1}{p}\sum_{i=1}^{p} y_{i\bullet\bullet}^2 - \frac{y_{\bullet\bullet\bullet}^2}{N}$$

$$= \frac{1}{5}\left[(-14)^2 + 9^2 + 5^2 + 3^2 + 7^2\right] - \frac{(10)^2}{25} = 68.00$$

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

$$SS_{Operations} = \frac{1}{p}\sum_{k=1}^{p} y_{\bullet\bullet k}^2 - \frac{y_{\bullet\bullet\bullet}^2}{N}$$

$$= \frac{1}{5}\left[(-18)^2 + 18^2 + (-4)^2 + 5^2 + 9^2\right] - \frac{(10)^2}{25} = 150.00$$

The total for the treatment are

Latin Letter	Treatment Total
A	$y_{\bullet 1\bullet} = 18$
B	$y_{\bullet 2\bullet} = -24$
C	$y_{\bullet 3\bullet} = -13$
D	$y_{\bullet 4\bullet} = 24$
E	$y_{\bullet 5\bullet} = 5$

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

The sum of square resulting from the formulations is computed from these totals as

$$SS_{Formulatios} = \frac{1}{p}\sum_{j=1}^{p} y_{\bullet j\bullet}^2 - \frac{y_{\bullet\bullet}^2}{N}$$

$$= \frac{(18)^2 + (-24)^2 + (-13)^2 + (24)^2 + (5)^2}{5} - \frac{(-10)}{25} = 330.00$$

The error of squares is found by subtraction

$$SS_E = SS_T - SS_{Batches} - SS_{Operators} - SS_{Formulations}$$

$$= 676.00 - 68.00 - 150.00 - 330.00 = 128.00$$

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

Analysis of variance for the Rocket Propellent Experiment

Sorces of variation	Sum of Squares	Degree of Freedom	Mean Square	F
Formation	330.00	4	82.25	7.73
Batches of raw material	68.00	4	17.00	
Operators	150.00	4	37.50	
Error	128.00	12	10.67	
Total	676.00	24		

Edited with the trial version of
Foxit Advanced PDF Editor
To remove this notice, visit:
www.foxitsoftware.com/shopping

Other Topics

- Missing values in blocked designs
  - RCBD
  - Latin square
- Replication of Latin Squares
- Crossover designs
- Graeco-Latin Squares
- Incomplete block designs